ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ
INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT




УДК 004.946
https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Научная статья



Визуальная когерентность в дополненной реальности

А.Л. Горбунов [1,2]
[1]Московский государственный технический университет гражданской авиации, г. Москва, Российская Федерация
[2]ООО «Авиареал», г. Москва, Российская Федерация
✉ a-gorbunov@mail.ru

Аннотация
Введение. 2020-е годы ознаменовались появлением нового поколения компьютерных тренажеров с применением технологии дополненной реальности. Одно из преимуществ данной технологии — возможность безопасного моделирования опасных ситуаций в реальном мире. Необходимым условием использования этого преимущества является обеспечение визуальной когерентности сцен дополненной реальности: виртуальные объекты должны быть неотличимы от реальных. Все мировые IT-лидеры рассматривают дополненную реальность как следующую волну радикальных изменений в цифровой среде, поэтому визуальная когерентность становится ключевым вопросом для будущего IT, а в аэрокосмических приложениях визуальная когерентность уже приобрела практическое значение. Примером может служить разработка корпорацией Боинг пилотского тренажера с дополненной реальностью (2022). Визуальная когерентность — сложная комплексная проблема, одним из аспектов которой является обеспечение корректной колористической тонировки виртуальных объектов в сцене дополненной реальности. Цель работы — разработка нового метода такой тонировки.

Материалы и методы. В разработанном методе (названном спектральной трансплантацией) используются двумерные спектральные преобразования изображений.

Результаты исследования. Предложена технология спектральной трансплантации, обеспечивающая прямую передачу характеристик цвета, яркости и контраста от реального фона к виртуальным объектам. Разработан алгоритм автоматического выбора оптимального вида спектрального преобразования.

Обсуждение и заключение. Будучи полностью автоматическим процессом без регистрации условий освещенности, спектральная трансплантация решает ряд сложных проблем визуальной когерентности. Спектральная трансплантация может стать ценным дополнением к другим методам обеспечения визуальной когерентности.

Ключевые слова: компьютерные тренажеры, дополненная реальность, визуальная когерентность

Благодарности: автор выражает благодарность А. Теренци (Inglobe Technologies Srl, Чеккано, Италия) за поддержку в разработке программного обеспечения.

Для цитирования. Горбунов А.Л. Визуальная когерентность в дополненной реальности. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2):180–190. https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Original article

Visual Coherence for Augmented Reality

Andrey L. Gorbunov [1,2]
[1] Moscow State Technical University of Civil Aviation, Moscow, Russian Federation
[2] "Aviareal" LLC, Moscow, Russian Federation
a-gorbunov@mail.ru

Abstract

Introduction. The 2020s were marked by the emergence of a new generation of computer simulators using augmented reality. One of the promising advantages of augmented reality technology is the ability to safely simulate hazardous situations real-world. A prerequisite for realizing this advantage is to provide the visual coherence of augmented reality scenes: virtual objects must be indistinguishable from real ones. All IT leaders consider augmented reality as a next "big wave"; thus, the visual coherence is becoming a key issue for IT in general. However, it is in aerospace applications that the visual coherence has already acquired practical significance. An example is Boeing's development of an augmented reality flight simulator, which began in 2022. Visual coherence is a complex problem, one of the aspects of which is to provide the correct overall coloration of virtual objects in an augmented reality scene. The objective of the research was to develop a new method of such tinting.

Materials and Methods. The developed method (called spectral transplantation) uses two-dimensional spectral image transformations.

Results. A spectral transplantation technology is proposed that provides direct transfer of color, brightness, and contrast characteristics from the real background to virtual objects. An algorithm for automatic selection of the optimal type of spectral transformation has been developed.

Discussion and Conclusion. Being a fully automatic process without recording lighting conditions, spectral transplantation solves a number of complex problems of visual coherence. Spectral transplantation can be a valuable addition to other methods of providing visual coherence.

Keywords: computer simulators, augmented reality, visual coherence

Acknowledgements: the author would like to thank Alessandro Terenzi, Inglobe Technologies Srl, Ceccano, Italy, for his support in software development.

For citation. Gorbunov AL. Visual Coherence for Augmented Reality. *Advanced Engineering Research (Rostov-on-Don)*. 2023;23(2): 180–190. https://doi.org/10.23947/2687-1653-2023-23-2-180-190

Введение. Современные тренажеры фактически по умолчанию подразумевают использование виртуальной реальности (virtual reality — VR). Преимущества такого подхода хорошо известны, поэтому не будем на них останавливаться, но отметим ряд существенных и, что важнее, непреодолимых недостатков, обусловленных самой природой технологии виртуальной реальности. VR — это цифровая, дискретная технология, в то время как реальный мир непрерывен. Поэтому моделирование реального мира в VR неизбежно связано с ошибками, что снижает эффективность обучения. Однако для обучающих систем еще более серьезным негативным аспектом является то, что решения человека во многом основаны на подсознательном учете множества деталей реальной картины мира. Этот процесс принципиально невозможно воспроизвести с помощью чисто компьютерных технологий (например, VR) по двум причинам: мы до сих пор не знаем (и вряд ли когда-либо узнаем) каков механизм работы человеческого мозга, последние спекуляции на тему искусственного интеллекта только подтверждают это; детали реального мира, учитываемые при принятии решений, практически бесконечны по количеству, возникают случайным образом и имеют самую разную природу (визуальную, акустическую, тактильную...).

Появление в 2020-х годах обучающих систем дополненной реальности (augmented reality — AR) снижает остроту этой проблемной ситуации. Примерами могут служить разработка компанией Boeing пилотского тренажера дополненной реальности на основе известного проекта R6 ATARS, которая началась осенью 2022 года, а также параллельно стартовавший похожий проект британской BAE Systems или тренажер авиадиспетчеров из данной статьи. Все информационное богатство окружающего нас мира в AR представлено в явном виде и не требует моделирования. Но для реализации преимуществ AR, связанных с параллельным присутствием реальных и виртуальных объектов в сценах, необходимо решить проблему визуальной когерентности (visual coherence — VC): виртуальные объекты должны быть неотличимы от реальных. В данной

статье предлагается метод решения проблемы визуальной когерентности в рамках проекта по разработке системы обучения авиадиспетчеров.

AR является производной формой VR. AR сохраняет все возможности VR, но, помимо этого, как гибридная технология обладает значительными преимуществами, вытекающими из параллельного сосуществования виртуальных и реальных объектов, что и привлекает внимание разработчиков к VC. Более того, исследования [1] показывают, что среди негативных психофизиологических последствий использования устройств дополненной реальности доминирует оптический дискомфорт, который возникает из-за разницы в восприятии реальных и виртуальных объектов в одной и той же сцене из-за отсутствия VC. Лидеры IT-отрасли рассматривают AR как следующую «большую волну» революционных изменений в цифровой электронике. Поэтому проблема VC становится ключевой для IT в целом, и эти лидеры проявляют растущий интерес к методам ее решения [2]. Однако в аэрокосмических приложениях проблема визуальной когерентности уже приобрела практическое значение. Авторы столкнулись с проблемой VC при разработке системы обучения авиадиспетчеров: стремительный рост интенсивности воздушного движения в аэропортах привел к увеличению частоты столкновений самолетов с другими воздушными судами и аэродромным транспортом при наземном маневрировании (>50 случаев по всему миру в 2018 году до начала пандемии). Авиадиспетчеры, работающие на вышках аэропортов, не всегда готовы адекватно реагировать на подобные нештатные ситуации, что требует дополнительной подготовки. Наиболее эффективная форма такого обучения предполагает представление диспетчеру ситуации опасного сближения объектов на летном поле, что невозможно с реальными объектами, но может быть абсолютно безопасно реализовано в сценах дополненной реальности. В нашем приложении аварийные ситуации были безопасно смоделированы с помощью AR на реальном аэродроме, при этом используемые виртуальные самолеты должны быть неотличимы от реальных.

Исчерпывающий обзор известных методов VC можно найти в статье [3]. Согласно авторской классификации, все методы VC можно разделить на два основных класса: с измерением параметров освещения и с оценкой условий освещения. В первом случае обязательной процедурой является предварительное измерение условий освещенности, проводимое с помощью специального оборудования. Эта процедура — длительный и трудоемкий процесс и не представляется возможной, если используется заранее полученное изображение или видео реального мира. Во втором случае сложность восстановления картины освещения по изображениям приводит к допущениям и ограничениям, что делает результаты неоднозначными. Поэтому, несмотря на впечатляющие результаты, полученные исследователями с помощью методов упомянутых в обзоре [3], уровень VC все еще часто недостаточен, особенно в сценах AR с реальными естественными ландшафтами в условиях рассеянного освещения, которые типичны для авиационных приложений. Как показывает приведенный ниже обзор публикаций, существует дефицит исследований такого рода.

Цель данной работы — разработка универсального и автоматического метода, обеспечивающего прямой перенос характеристик цвета, яркости и контраста с реального фона на виртуальные объекты без цифрового 3D-моделирования, которое необходимо в существующих подходах VC. Метод основан на математическом аппарате двумерных спектральных преобразований, мы назвали его «спектральной трансплантацией».

Основными результатами данной работы являются:

– базовая схема для метода спектральной трансплантации, который обеспечивает прямой перенос характеристик цвета, яркости и контраста с реального фона на виртуальные объекты. Метод предполагает замену части спектра изображения виртуального мира на такую же часть спектра изображения реального мира с последующим обратным преобразованием спектра с пересаженной частью;

– алгоритм автоматического выбора оптимального вида спектрального преобразования для использования в спектральной трансплантации.

Важно отметить, что VC зависит от многих факторов: освещения, теней, цветового тона, взаимных отражений, текстуры поверхности, оптических аберраций, конвергенции, аккомодации и пр. Соответственно используются различные техники AR-визуализации. В нашем случае VC обеспечивается только для факторов общего освещения и окраски виртуальных объектов в AR. Это является одним из наиболее сложных вопросов VC, особенно для сцен на открытом воздухе. Поэтому для достижения полной VC спектральная трансплантация должна использоваться в сочетании с другими методами VC.

Список источников в [3] включает 175 позиций; этот обзор включает практически все подходы к достижениям в VC (за исключением последних, основанных на нейронных сетях, рассматриваемых ниже). Поэтому здесь кратко опишем некоторые характерные примеры, которые соответствуют упомянутым основным классам.

Измерение условий освещения

Используя зонд освещенности с диффузными полосами между зеркальными сферическими квадрантами, Дебевек и др. [4] продемонстрировали, как полный динамический цветовой диапазон сцены может быть восстановлен из одной экспозиции. На основе изображения, полученного с помощью зонда, интенсивность нескольких источников света может быть оценена путем решения простой линейной системы уравнений. Результаты были использованы для рендеринга виртуальной диффузной сферы.

Альхаками и Тасерян [5] оценивают направление падающего света (прямое освещение) реальной сцены используя методы компьютерного зрения с помощью 360°- камеры, прикрепленной к AR-устройству. Система моделирует отраженный свет от поверхностей при рендеринге виртуальных объектов. Затем определяются параметры теней для каждого виртуального объекта.

Оценка условий освещения

Кнорр и Курц [6] предложили схему оценки условий освещения в реальном мире на основе фото человеческого лица. Метод основан на обучении модели вида лица на основе базы данных по лицам при известном освещении. Авторы затем восстанавливают наиболее правдоподобные условия освещения в реальном мире в базисе сферических гармоник для запечатленного лица.

Стоит упомянуть работу [7], где описана комбинация измерения и оценки освещенности. Авторы измеряют отражательные свойства реальных объектов, используя карты глубины и цветные изображения вращающегося объекта на поворотном столе с помощью RGB-D камеры. Форма объекта реконструируется путем интеграции изображений глубины объекта, полученных с разных точек обзора. Отражательная способность объекта определяется путем оценки параметров модели отражения по восстановленным изображениям формы и цвета.

Наиболее близкими аналогами предлагаемого метода являются подходы, которые, как и спектральная трансплантация не предполагают предварительных измерений освещения и моделирования условий освещения, геометрии сцены, отражения поверхности, а также предусматривают автоматическую обработку.

Среди таких аналогов — методы передачи цвета от изображения к изображению. В работе [8] представлен способ автоматического переноса цветовой статистики (средних значений и стандартных отклонений) с эталонного изображения на целевое. Дополнительные параметры используются для того, чтобы избежать ручной обработки, которая необходима для определения особенностей передачи цвета в случаях, когда изображения имеют сильное различие в цветовой палитре. Эти дополнительные параметры объединяют дисперсии эталонного и целевого изображений. Авторы статьи утверждают, что, хотя ручное изменение этих параметров требуется крайне редко, тем не менее иногда это необходимо. Кроме того, статистическая природа метода вызывает вопросы относительно вида и объема статистики. Также неочевидна способность метода обрабатывать некоторые типы изображений (содержащие блестящие объекты, тени).

Сяо и Ма [9] представили алгоритм для решения проблемы достоверности передачи цвета с точки зрения деталей сцены и цветов. Авторы рассматривают сохранение цветового градиента как необходимое условие достоверности сцены. Они формулируют проблему передачи цвета как задачу оптимизации и решают ее в два этапа — согласование гистограмм и оптимизация с сохранением градиента. Предложена метрика для объективной оценки эффективности алгоритмов передачи цвета на основе примеров.

Преимуществами разработанного метода, по сравнению с [8, 9] и их многочисленными аналогами, являются его универсальность, полностью автоматический характер и возможность передачи не только цвета, но и всех основных характеристик изображения с помощью одной простой процедуры.

Предложенный метод использует двумерные спектральные преобразования. Различные виды изображений оптимально описываются различными типами спектральных преобразований («оптимально» — в смысле соответствия визуального восприятия для реальных и виртуальных объектов). Среди активно используемых в цифровой обработке изображений со времен появления цифрового телевидения — дискретное преобразование Фурье, дискретное косинусное преобразование, преобразование Хадамара, S-преобразование и преобразование Карунена-Лоэва.

Материалы и методы. Схема метода спектральной трансплантации (версия, использующая преобразование Фурье [10]) показана на рис. 1. В качестве входных данных используются кадры реального мира (world frame — WF) и виртуального мира (virtual frame — VF) (рис. 2). Это естественно для «видео» AR (когда реальный мир наблюдается через видеокамеру). Для «прозрачной» AR, когда реальный мир наблюдается через прозрачные очки, реальные изображения захватываются с помощью камер, расположенных на AR-очках.

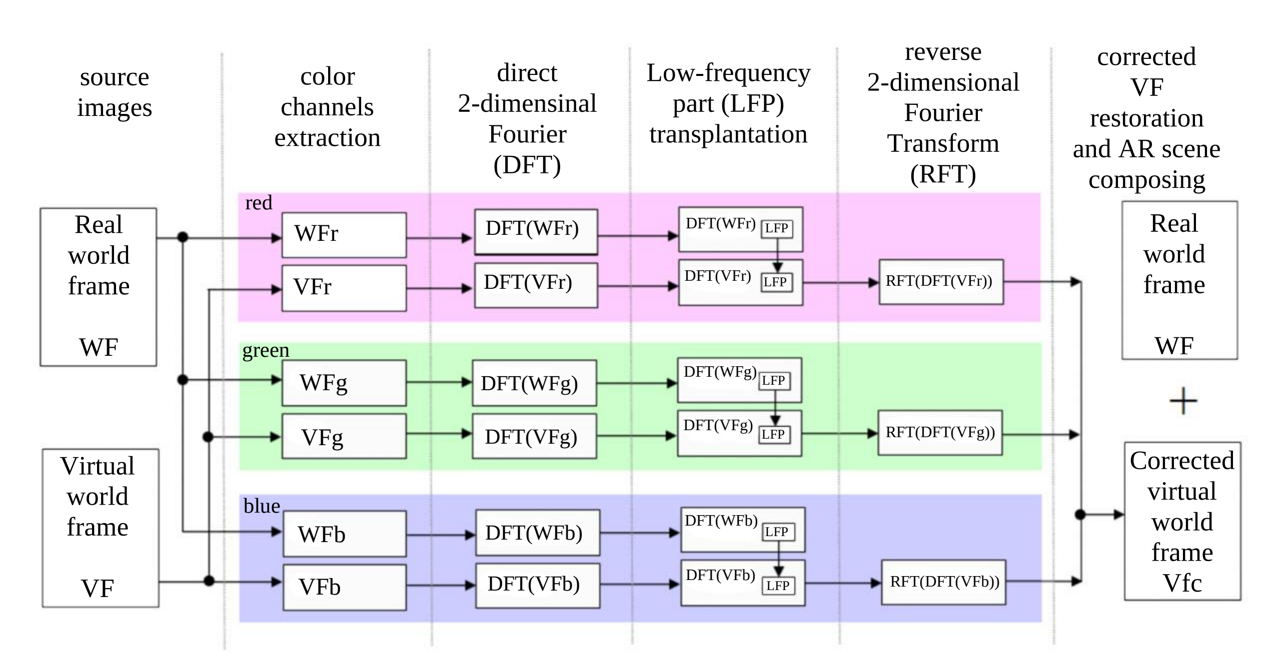
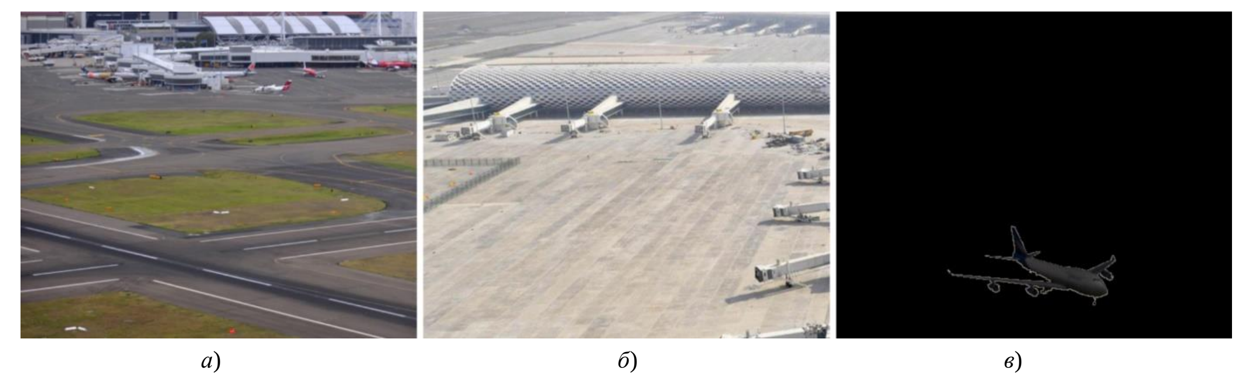


Рис. 1. Схема метода спектральной трансплантации.
Версия, использующая преобразование Фурье



а) *б*) *в*)

Рис. 2. Видеокадры реального мира (WF) и виртуального мира (VF):
а — WF, аэропорт, облачная погода; *б* — WF, аэропорт, солнечная погода; *в* — VF, виртуальный самолет. WF представляют собой небольшие фрагменты (<25 %) изображений, опубликованных на сайтах sydneyairport.com.au и 6sqft.com

Целью данного метода является перенос основных характеристик изображения из WF в VF. Схема работы метода очень проста, хотя операции имеют большой вычислительный объем. Метод реализуется в пять этапов (рис. 1):

1) Выделение цветовых (RGB) каналов для WF — WFr, WFg, WFb и для VF — VFr, VFg, VFb. Модель RGB используется из-за ее общности и корреляции между каналами, которая специфична для этой модели.

2) Вычисление двумерного прямого преобразования Фурье (direct Fourier transform — DFT): DFT(WFr), DFT(WFg), DFT(WFb), DFT(VFr), DFT(VFg), DFT(VFb). Формула DFT приведена ниже:

$$X_c(k,l) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x_c(m,n)\exp(-j2\pi(\frac{km}{M}+\frac{nl}{N})) \quad (1)$$

где $c = R, G, B$ — индекс для красного, зеленого и синего каналов цветного изображения; M, N — номера строк и столбцов матрицы пикселей преобразованного изображения; k, l — аргументы пространственной частоты; $x_c(m,n)$ — значение пикселя с пространственными координатами (m,n) в канале c; $X_c(k,l)$ — комплексные числа.

3) Этот этап является ключевым. Трансплантация низкочастотной части (low-frequency part-LFP) осуществляется между парами спектров WF и VF для каждого из красного, зеленого и синего каналов. Это означает, что LFP VF заменяется на соответствующую LFP WF. Идея спектральной трансплантации основана на следующем свойстве DFT: общий характер изображения (т.е. цветовой оттенок, яркость, контраст) зависит от пространственных частот, содержащихся в LFP (включая постоянную составляющую) его двумерного спектра.

Таким образом, пересаживая LFP WF в спектр VF, мы переносим основные характеристики изображения из WF в VF. Для этого удобнее использовать центрированную форму двумерного спектра, где постоянная составляющая находится в центре матрицы спектральных коэффициентов, а низкочастотные компоненты симметрично расположены вокруг постоянной составляющей. В центрированном спектре LFP является центральной частью матрицы DFT, и LFP имеет размер Ml × Nl (Ml <M, Nl <N). Если Ml =Nl, то обозначение для квадратной матрицы LFP может быть LFP(012..F), где 0 — постоянная компонента, F — номер наибольшей пространственной частоты в матрице LFP.

Размер LFP для трансплантации зависит от размера преобразованного изображения (этот размер определяет спектральное разрешение) и от объема характеристик изображения, которые должны быть заимствованы из WF. На данном этапе исследований размер LFP определяется эмпирически. Например, наилучшие визуальные результаты для изображений размером 512×512 пикселей были получены при использовании LFP (012345).

4) Восстановление каналов RGB для VF с помощью двумерного обратного преобразования Фурье (reverse Fourier transform — RFT). Одновременно на этом этапе происходит смешение характеристик WF и VF. В результате получаются RGB-каналы изображения VF с основными цветовыми, яркостными и контрастными характеристиками WF, а также с характеристиками, унаследованными от исходного VF.

5) Восстановление откорректированного цвета VF путем слияния RGB-каналов, полученных на предыдущем этапе, вырезания виртуальных объектов и построения AR-сцены путем наложения вырезанных виртуальных объектов на WF.

Очевидно, что если этот метод используется для обработки видеопотока WF, то нет необходимости в вычислениях DFT и, соответственно, LFP для каждого кадра реального мира, поскольку основные характеристики изображения меняются только при радикальном изменении записываемой сцены. Такие изменения можно легко обнаружить по скачкам среднего значения пикселей. В эти моменты необходимо пересчитать спектральное преобразование для LFP.

Поскольку различные виды изображений оптимально описываются различными видами спектральных преобразований упомянутыми выше, целесообразно разработать автоматический алгоритм выбора оптимального вида преобразования для использования в спектральной трансплантации.

Мы предлагаем оценивать разницу между зрительным восприятием VF и WF по среднеквадратичному расстоянию *Δ* между LFP спектров мощности их изображений (для всех цветовых каналов):

$$\Delta_c = \frac{1}{M_l N_l}\sum_{k=0}^{M_l}\sum_{l=0}^{N_l}\left|P_{Vc}(k,l) - P_{Wc}(k,l)\right|^2, \; c = R, G, B, \tag{2}$$

где P_V и P_W — двумерные спектры мощности соответственно VF и WF. Например, в случае преобразования Фурье, формула для P имеет вид:

$$P_c(k,l) = \left|\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x_c(m,n)\exp(-j2\pi(\frac{km}{M}+\frac{nl}{N}))\right|^2. \tag{3}$$

Мы предлагаем определять оптимальный вид спектрального преобразования близостью векторов *Δ* и среднего вектора, вычисленного по критерию минимальной суммы квадратов расстояний между средним вектором и векторами *Δ* для всех рассматриваемых преобразований.

Пусть $\Delta_j(\Delta_j R, \Delta_j G, \Delta_j B)$ — нормированный вектор расстояния между спектром LFPs VF и WF для преобразования j. Пусть $\Delta_a(\Delta_a R, \Delta_a G, \Delta_a B)$ — средний вектор, а D_j — расстояние между Δ_j и Δ_a. Тогда сумма S квадратов расстояний от векторов Δ_j всех рассматриваемых преобразований до среднего вектора равна:

$$S = \sum_j {D_j}^2 = \sum_j\sum_c(\Delta_{jc} - \Delta_{ac})^2, \; c = R, G, B. \tag{4}$$

Координаты Δ_{aR}, Δ_{aG}, Δ_{aB} среднего вектора вычисляются как решение системы дифференциальных уравнений с частными производными:

$$\frac{\partial S}{\partial \Delta_{ac}} = 0, \; c = R, G, B. \tag{5}$$

Выбор оптимального вида спектрального преобразования определяется условием близости:

$$\min_j D_j. \tag{6}$$

Другим очевидным критерием выбора оптимального вида преобразования является длина векторов *Δ*. Однако экстремумы такого критерия могут быть связаны со способностью или неспособностью определенных преобразований правильно обнаружить разницу между определенными видами WF и VF. Поэтому мы рассматриваем использование среднего вектора как более надежный способ отбора.

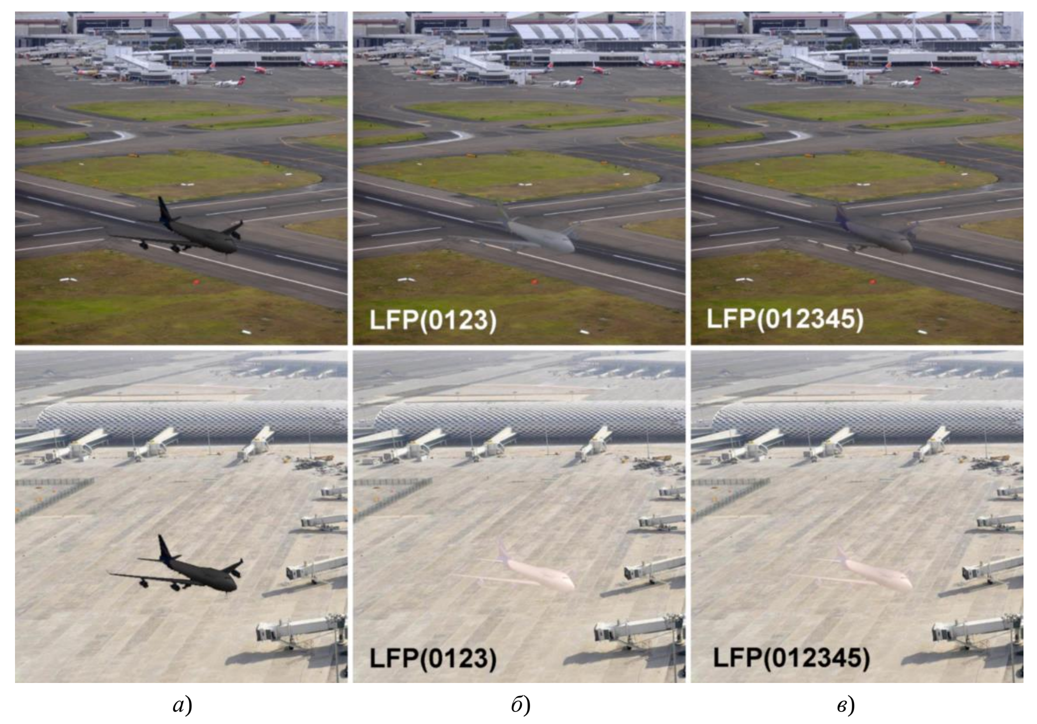
Аналогично расчету DFT для WF, выбор оптимального вида преобразования производится только один раз в начале спектральной трансплантации, если только WF не меняется резко.

Результаты исследования. Предложенный метод был протестирован с использованием преобразования Фурье без выбора оптимального вида преобразования. WF (реальная сцена аэропорта) и VF (виртуальная модель самолета) имели размер 512×12пикселей и 24-битные цвета. Исследовали два различных условия:

1) WF — фото аэропорта в облачную погоду, рис. 2 *а*;

2) WF — фото аэропорта в солнечную погоду, рис. 2 *б*.

В обоих случаях VF содержал 3D-модель самолета, показанную на рис. 2 *в*. Протестировали трансплантации LFP(0), LFP(01), LFP(012), LFP(0123), LFP(01234), LFP(012345). Некоторые результаты тестирования показаны на рис. 3. Наилучшие визуальные результаты были получены при использовании LFP(012345). На рис. 3 изображения после спектральной трансплантации намеренно показаны без других эффектов VC (тени, освещение и т.д.), чтобы продемонстрировать чистые результаты этого метода.



а) *б*) *в*)

Рис. 3. AR-сцены без трансплантации и с трансплантацией LFP разных размеров.
а — AR-сцена, состоящая из WF и VF без трансплантации LFP; *б* — AP-сцена, после трансплантации LFP(0123); *в* — AR-сцена после трансплантации LFP(012345)

Верхний и нижний ряды на рис. 3 соответствуют противоположным условиям для WF: светлые и темные WF с разными оттенками. Эксперименты с любыми другими WF не добавят существенно новой информации, поскольку будут иметь условия между теми, что уже представлены на рис. 3.

Для демонстрации механизма спектральной трансплантации было проведено численное моделирование. На рис. 4 показана Фурье-трансплантация с использованием небольшой (8×8) пиксельной матрицы, представляющей один из цветовых каналов WF и VF. Такой маленький размер матрицы позволяет наглядно проиллюстрировать процедуру трансплантации. В данном примере матрица WF может быть ассоциирована с изображением с вертикальной градиентной заливкой, а матрица VF — с изображением с горизонтальной градиентной заливкой. Другим отличием между WF и VF является диапазон значений пикселей: 8–15 для WF («более светлое изображение», 8 — постоянная составляющая) и 0–7 для VF («более темное изображение»). Показана трансплантация LFP(01'), где 1' означает часть первой пространственной частотной компоненты (используется из-за очень низкого разрешения матрицы 8×8). 3D вид матрицы ВФ после трансплантации свидетельствует о передаче свойств от WF к VF: край поверхности приподнялся, первый пиксель получил значение постоянной составляющей WF. Этот пример демонстрирует, как в результате спектральной трансплантации VF начинает приобретать вертикальный градиент и постоянную составляющую.

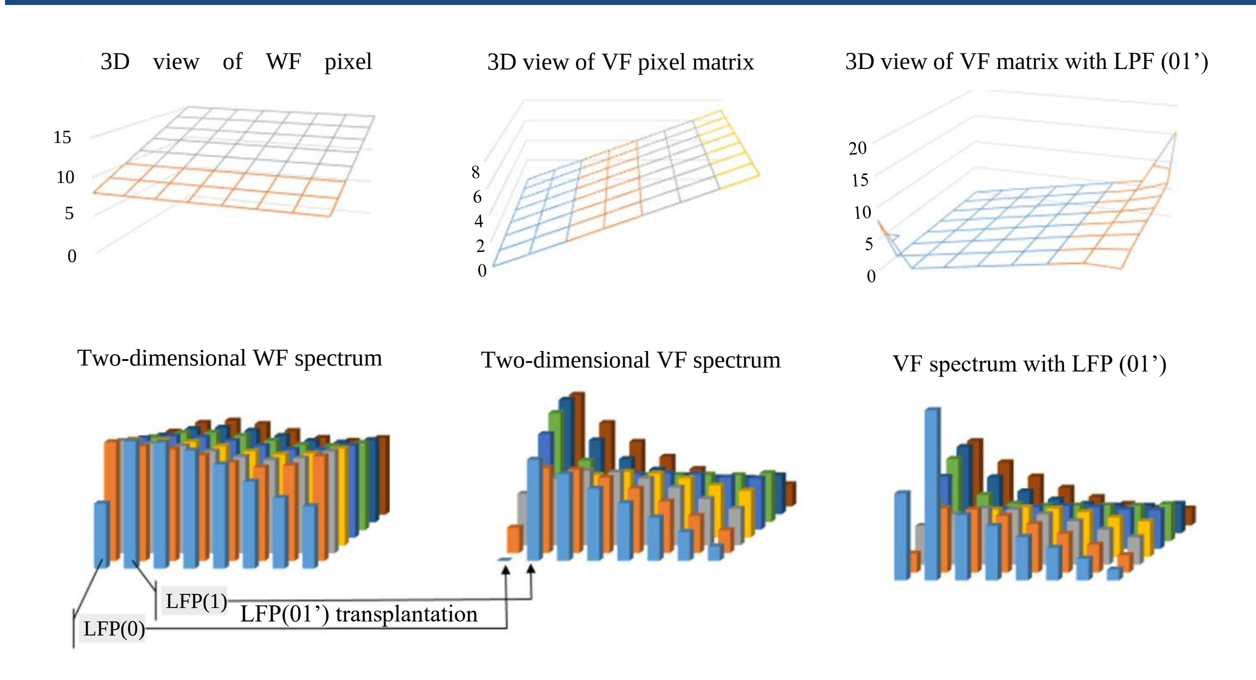


Рис. 4. Численное моделирование спектральной трансплантации для матриц 8×8 пикселей

Спектральная трансплантация предоставляет несколько вариантов изменения параметров этой процедуры: изменение размера LFP; выбор отдельных компонентов спектра для трансплантации; использование различных коэффициентов трансплантации для разных трансплантируемых компонентов.

На рис. 5 показан эффект трансплантации с различными параметрами для различных типов виртуальных объектов — виртуальных моделей самолетов, отличающихся текстурой поверхности, маркировкой и блеском. На рис. 5 *а* и *б* изображен виртуальный самолет со сложными текстурами, текстовыми обозначениями и отражениями виртуальных источников света. На рис. 5 *г*, д и *е* — виртуальный самолет с простыми контрастными цветами. Части *а* и *г* содержат виртуальные объекты без трансплантации; *б* и *д* содержат виртуальные объекты после трансплантации LFP(0123); *в* и *е* содержат виртуальные объекты после трансплантации LFP(012345). Виртуальные объекты намеренно показаны без других эффектов VC (тени, освещение и т.д.), чтобы продемонстрировать чистые результаты метода.



Рис. 5. Сцены с пасмурной WF: *а, г* — AR-сцены, состоящие из WF и VF без трансплантации LFP; *б, д* — AP-сцены, составленные после трансплантации LFP(0123); *в, е* — AR-сцены, составленные после трансплантации LFP(012345)

Важно подчеркнуть, что представленные рисунки иллюстрируют возможности настройки предложенного метода, а не конечный результат, поскольку он требует настройки на конкретный WF. Демонстрация хорошо выполненных, но неполных результатов, как это часто практикуется в работах по VC, не кажется нам корректной.

Обсуждение и заключение. Основным осложняющим фактором для описанного метода, представленного на рис. 1, являются высокие вычислительные затраты. Наиболее перспективным способом решения этой проблемы является прямое преобразование параметров LFP WF в параметры рендеринга VF. Это позволяет отказаться от громоздких процедур трех DFT- и трех RFT-расчетов на втором и четвертом шагах обработки и требует только трех расчетов WF DFT, по одному разу для каждого участка потока WF без существенного изменения базовых характеристик WF. Такой подход позволяет обрабатывать поток VF в режиме реального времени.

Другой проблемой является выбор оптимального размера LFP. По мере увеличения объема используемых пространственных частот, они начинают содержать информацию о содержимом WF, поэтому ограничение размера LFP необходимо для устранения эффекта гибридного изображения [11]. Сложность оптимального выбора обусловлена тем, что он связан как с размером LFP, так и с характером изображения. Последние достижения в области глубокого обучения позволяют предположить, что новым подходом, связанным с визуальной когерентностью посредством спектральной трансплантации, может стать использование генеративно-состязательных сетей (generative adversarial network — GAN) для передачи реалистичной информации об освещении с исходного изображения на целевое изображение тем же образом, как это делают GAN для передачи стиля изображения В частности, было бы интересно сравнить работу GAN в случае наборов данных, состоящих либо из RGB-изображений, либо из изображений, представленных в частотной области с помощью DFT. Мы считаем, что использование последнего подхода поможет выбрать оптимальный LFP. GAN уже широко используются в исследованиях VC [2], как и нейронные сети в целом [12].

В дальнейших исследованиях, связанных с темой данной работы, будут рассмотрены следующие вопросы:

– автоматическое определение оптимального размера LFP для трансплантации с заданным объемом характеристик, заимствуемых из WF;

– автоматическое определение точных моментов, когда необходимо рассчитать новую LFP WP для трансплантации при обработке видео WF в реальном времени (как указывалось выше, это необходимо делать, если базовые характеристики WF радикально меняются);

– использование того же подхода в обратном направлении (от виртуального к реальному) для применения виртуального освещения к реальным сценам (как виртуальное освещение влияет на окружающую среду).

Как полностью автоматический процесс без измерения освещения, предлагаемый метод спектральной трансплантации решает ряд сложных проблем VC. Например, как наилучшим образом выровнять характеристики цвета, яркости и контраста между реальными и виртуальными компонентами в сценах AR. Все эти задачи решаются посредством одной простой процедуры без моделирования условий освещения, геометрии AR-сцены или BRDF, что исключает неизбежные ошибки моделирования. Предложенный метод может стать ценным дополнением к другим инструментам VC.

Список литературы

1. Hughes C.L., Fidopiastis C., Stanney K.M., et al. The Psychometrics of Cybersickness in Augmented Reality. *Frontiers in Virtual Reality*. 2020;1:602954. https://doi.org/10.3389/frvir.2020.602954

2. Somanath G., Kurz D. *HDR Environment Map Estimation for Real-Time Augmented Reality*. Cupertino, CA: Apple Inc.; 2020. URL: https://arxiv.org/pdf/2011.10687.pdf (accessed: 17.06.2022).

3. Kronander J., Banterle F., Gardner A., et al. Photorealistic Rendering of Mixed Reality Scenes. *Computer Graphics Forum*. 2015;34(2):643–665. https://doi.org/10.1111/cgf.12591

4. Debevec P., Graham P., Busch J., et al. A Single-Shot Light Probe. *SIGGRAPH '12: ACM SIGGRAPH 2012 Talks*. 2012;10:1–19. URL: https://vgl.ict.usc.edu/Research/SSLP/A_Single_Shot_Light_Probe-SIGGRAPH2012.pdf (accessed: 17.06.2022).

5. Alhakamy A., Tuceryan M. *CubeMap360: Interactive Global Illumination for Augmented Reality in Dynamic Environment*. In: Proc. IEEE SoutheastCon. Huntsville, AL: IEEE; 2019. https://doi.org/10.1109/SoutheastCon42311.2019.9020588

6. Knorr S.B., Kurz D. *Real-Time Illumination Estimation from Faces for Coherent Rendering*. In: Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR). Munich: IEEE; 2014. P. 113–122. https://doi.org/10.1109/ISMAR.2014.6948483

7. Seiji Tsunezaki, Ryota Nomura, Takashi Komuro, et al. *Reproducing Material Appearance of Real Objects using Mobile Augmented Reality*. In: Proc. 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Munich: IEEE; 2018. P. 196–197. https://doi.org/10.1109/ISMAR-Adjunct.2018.00065

8. Reinhard E., Akyuz A.O., Colbert M., et al. *Real-Time Color Blending of Rendered and Captured Video*. In: Proc. Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). Orlando, Florida: National Training and Simulation Association; 2004. P. 1–9. URL: https://user.ceng.metu.edu.tr/~akyuz/files/blend.pdf (accessed: 17.06.2022).

9. Xuezhong Xiao, Lizhuang Ma. Gradient-Preserving Color Transfer. *Computer Graphics Forum*. 2009;28(7):1879–1886. https://doi.org/10.1111/j.1467-8659.2009.01566.x

10. Горбунов А.Л. и др. *Способ формирования изображения дополненной реальности, обеспечивающий совпадение визуальных характеристик реальных и виртуальных объектов*. Патент РФ, № 2667602, 2019.

11. Oliva A., Torralba A.J., Schyns Ph.G. Hybrid Images. *ACM Transactions on Graphics*. 2006;25(3):527–532. https://doi.org/10.1145/1179352.1141919

12. Kán P., Kafumann H. DeepLight: Light Source Estimation for Augmented Reality Using Deep Learning. *The Visual Computer*. 2019;35:873–883. https://doi.org/10.1007/s00371-019-01666-x

References

1. Hughes CL, Fidopiastis C, Stanney KM, et al. The Psychometrics of Cybersickness in Augmented Reality. *Frontiers in Virtual Reality*. 2020;1:602954. https://doi.org/10.3389/frvir.2020.602954

2. Somanath G, Kurz D. *HDR Environment Map Estimation for Real-Time Augmented Reality*. Cupertino, CA: Apple Inc.; 2020. URL: https://arxiv.org/pdf/2011.10687.pdf (accessed: 17.06.2022).

3. Kronander J, Banterle F, Gardner A, et al. Photorealistic Rendering of Mixed Reality Scenes. *Computer Graphics Forum*. 2015;34(2):643–665. https://doi.org/10.1111/cgf.12591 (accessed: 17.06.2022).

4. Debevec P, Graham P, Busch J, et al. A Single-Shot Light Probe. *SIGGRAPH '12: ACM SIGGRAPH 2012 Talks*. 2012;10:1–19. URL: https://vgl.ict.usc.edu/Research/SSLP/A_Single_Shot_Light_Probe-SIGGRAPH2012.pdf (accessed: 17.06.2022).

5. Alhakamy A, Tuceryan M. *CubeMap360: Interactive Global Illumination for Augmented Reality in Dynamic Environment*. In: Proc. IEEE SoutheastCon. Huntsville, AL: IEEE; 2019. https://doi.org/10.1109/SoutheastCon42311.2019.9020588

6. Knorr SB, Kurz D. *Real-Time Illumination Estimation from Faces for Coherent Rendering*. In: Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR). Munich: IEEE; 2014. P. 113–122. https://doi.org/10.1109/ISMAR.2014.6948483

7. Seiji Tsunezaki, Ryota Nomura, Takashi Komuro, et al. *Reproducing Material Appearance of Real Objects using Mobile Augmented Reality*. In: Proc. 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Munich: IEEE; 2018. P. 196–197. https://doi.org/10.1109/ISMAR-Adjunct.2018.00065

8. Reinhard E, Akyuz AO, Colbert M, et al. *Real-Time Color Blending of Rendered and Captured Video*. In: Proc. Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). Orlando, Florida: National Training and Simulation Association; 2004. P. 1–9. URL: https://user.ceng.metu.edu.tr/~akyuz/files/blend.pdf (accessed: 17.06.2022).

9. Xuezhong Xiao, Lizhuang Ma. Gradient-Preserving Color Transfer. *Computer Graphics Forum*. 2009;28(7):1879–1886. https://doi.org/10.1111/j.1467-8659.2009.01566.x

10. Gorbunov AL, et al. *Sposob formirovaniya izobrazheniya dopolnennoi real'nosti, obespechivayushchii sovpadenie vizual'nykh kharakteristik real'nykh i virtual'nykh ob"ektov*. RF Patent No. 2667602. 2019. (In Russ.)

11. Oliva A, Torralba AJ, Schyns PhG. Hybrid Images. *ACM Transactions on Graphics*. 2006;25(3):527–532. https://doi.org/10.1145/1179352.1141919

12. Kán P, Kafumann H. DeepLight: Light Source Estimation for Augmented Reality Using Deep Learning. *The Visual Computer*. 2019;35:873–883. https://doi.org/10.1007/s00371-019-01666-x

Поступила в редакцию 10.04.2023
Поступила после рецензирования 18.04.2023
Принята к публикации 08.04.2023

Об авторе:

Андрей Леонидович Горбунов, кандидат технических наук, доцент кафедры управления воздушным движением Московского государственного технического университета гражданской авиации (125993, РФ, г. Москва, Кронштадтский бул., 20), гендиректор ООО «Авиареал» (115372, РФ, г. Москва, ул. Загорьевская, 5), ORCID, ResearcherID, ScopusID, AuthorID, a-gorbunov@mail.ru

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Автор прочитал и одобрил окончательный вариант рукописи.

Received 10.04.2023
Revised 18.04.2023
Accepted 08.04.2023

About the Author:

Andrey L. Gorbunov, Cand.Sci. (Eng.), associate professor, Air Traffic Control Department, Moscow State Technical University of Civil Aviation (20, Kronshtadtskii blvd, Moscow, 125993, RF), Director-General, “Aviareal” LLC (5, Zagoryevskaya St., Moscow, 115372, RF), ORCID, ResearcherID, ScopusID, AuthorID, a-gorbunov@mail.ru

Conflict of interest statement: the author does not have any conflict of interest.

The author has read and approved the final manuscript.